



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/517,324

Source: PCT

Date Processed by STIC: 12-28-04

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

**FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.2 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

**<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

**Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:**

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>) , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):  
U.S. Patent and Trademark Office, 220 20<sup>th</sup> Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby, Room 1B03, Arlington, VA 22202

Revised 05/17/04



PCT

## RAW SEQUENCE LISTING

DATE: 12/28/2004

PATENT APPLICATION: US/10/517,324

TIME: 11:50:05

Input Set : A:\040679.txt

Output Set: N:\CRF4\12282004\J517324.raw

3 <110> APPLICANT: SinoGenoMax Company Ltd  
 5 <120> TITLE OF INVENTION: Randomised DNA libraries and double-stranded RNA libraries,  
 use and

6 method of production thereof

8 &lt;130&gt; FILE REFERENCE: P06031PC00

C--&gt; 10 &lt;140&gt; CURRENT APPLICATION NUMBER: US/10/517,324

C--&gt; 11 &lt;141&gt; CURRENT FILING DATE: 2004-12-20

13 &lt;150&gt; PRIOR APPLICATION NUMBER: US 60/390,108

14 &lt;151&gt; PRIOR FILING DATE: 2002-06-21

16 &lt;160&gt; NUMBER OF SEQ ID NOS: 45

18 &lt;170&gt; SOFTWARE: PatentIn version 3.1

20 &lt;210&gt; SEQ ID NO: 1

21 &lt;211&gt; LENGTH: 29

22 &lt;212&gt; TYPE: DNA

23 &lt;213&gt; ORGANISM: Artificial Sequence

25 &lt;220&gt; FEATURE:

26 &lt;223&gt; OTHER INFORMATION: 5?primer

28 &lt;400&gt; SEQUENCE: 1

29 ggaattcgaa cgctgacgac atcaaccgcg

32 &lt;210&gt; SEQ ID NO: 2

33 &lt;211&gt; LENGTH: 35

34 &lt;212&gt; TYPE: DNA

35 &lt;213&gt; ORGANISM: Artificial Sequence

37 &lt;220&gt; FEATURE:

38 &lt;223&gt; OTHER INFORMATION: 3?primer

40 &lt;400&gt; SEQUENCE: 2

41 gaagatctgt ctcatacaga acttataaga ttccc

44 &lt;210&gt; SEQ ID NO: 3

45 &lt;211&gt; LENGTH: 500

46 &lt;212&gt; TYPE: DNA

47 &lt;213&gt; ORGANISM: Artificial Sequence

49 &lt;220&gt; FEATURE:

50 &lt;223&gt; OTHER INFORMATION: Modified sequence

52 &lt;220&gt; FEATURE:

53 &lt;221&gt; NAME/KEY: misc\_feature

54 &lt;222&gt; LOCATION: (7)..(7)

55 <223> OTHER INFORMATION: Represents nucleotides not clearly identified in the  
 sequencing r

56 result.

59 &lt;220&gt; FEATURE:

60 &lt;221&gt; NAME/KEY: misc\_feature

61 &lt;222&gt; LOCATION: (9)..(9)

62 <223> OTHER INFORMATION: Represents nucleotides not clearly identified in the  
 sequencing r

↑ "N" can only represent a single nucleotide  
 file://C:\CRF4\OUTHOLD\VSrJ517324.htm

Does Not Comply

Corrupted Diskette Name

(pgs 1-4, 6)

Pls explain source of  
 genetic material.

Invalid  
 response

29

Pls explain source of genetic  
 material.

Invalid  
 response

35

Pls explain source of genetic  
 material.

Invalid response

See  
 iter  
 #11

Invalid response

On pr  
 sum  
 ste

12/28/04

63

66 <400> SEQUENCE: 3

result

part of  
- see error explanation on ps. 1

DATE: 12/28/2004

PATENT APPLICATION: US/10/517,324

TIME: 11:50:05

Input Set : A:\040679.txt

Output Set: N:\CRF4\12282004\J517324.raw

W--> 67 tccaggaac cggggcccag tgtcactagg cgggaacacc cagcgccgct gcgccctggc 60  
69 aggaagatgg ctgtgagggg caggggagtg gcgccctgca atatttgcgcat gtcgctatgt 120  
71 gttctgggaa atcaccataa acgtgaaatg tctttggatt tgggaatctt ataagttctg 180  
73 tatgagacag atcttcaata ttggccatta gccatattat tcattgggta tatagcataa 240  
75 atcaatattg gctattggcc attgcatacg ttgtatctat atcataatat gtacatttat 300  
77 attggctcat gtccaatatg accgccatgt tggcattgat tattgactag ttattaatat 360  
79 taatcaatta cgggggtcatt agttcatagc ccattatggg agttccgctg tacataaactt 420  
81 acggtaaatg gccgcctgg ctgaccgccc aacgaccccc gccattgac gtcaataatg 480  
83 acgtatgttc ccatagtaac 500  
86 <210> SEQ ID NO: 4  
87 <211> LENGTH: 38  
88 <212> TYPE: DNA  
89 <213> ORGANISM: Artificial Sequence  
91 <220> FEATURE:  
92 <223> OTHER INFORMATION: 5?rimer Same error  
94 <400> SEQUENCE: 4  
95 acgcgtcgac gaattcgaac gctgacgtca tcaaccgc 38  
98 <210> SEQ ID NO: 5  
99 <211> LENGTH: 36  
100 <212> TYPE: DNA  
101 <213> ORGANISM: Artificial Sequence  
103 <220> FEATURE:  
104 <223> OTHER INFORMATION: 3?rimer Same error  
106 <400> SEQUENCE: 5  
107 cccaagcttg tctcatagc aacttataag attccc 36  
110 <210> SEQ ID NO: 6  
111 <211> LENGTH: 50  
112 <212> TYPE: DNA  
113 <213> ORGANISM: Artificial Sequence  
115 <220> FEATURE:  
116 <223> OTHER INFORMATION: oligo DNA  
118 <400> SEQUENCE: 6  
119 ggggaagatc taaaaaata aatgaatcaa gaacattttt aagcttgggg 50  
122 <210> SEQ ID NO: 7  
123 <211> LENGTH: 50  
124 <212> TYPE: DNA  
125 <213> ORGANISM: Artificial Sequence  
127 <220> FEATURE:  
128 <223> OTHER INFORMATION: oligo DNA  
130 <400> SEQUENCE: 7  
131 cccaagctt aaaaatgttc ttgattcatt tattttttta gatcttcccc 50  
134 <210> SEQ ID NO: 8  
135 <211> LENGTH: 50  
136 <212> TYPE: DNA  
137 <213> ORGANISM: Artificial Sequence  
139 <220> FEATURE:  
140 <223> OTHER INFORMATION: 19 nucleotides randomized region  
142 <220> FEATURE:  
143 <221> NAME/KEY: misc feature  
Invalid

What is the source of genetic material?

Invalid response

↑ see item # 11 on

Dr. M. S. LAMM <sup>cr</sup> / 12/28/04

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/517,324

DATE: 12/28/2004

TIME: 11:50:05

Input Set : A:\040679.txt

Output Set: N:\CRF4\12282004\J517324.raw

144 <222> LOCATION: (17)..(35)  
 145 <223> OTHER INFORMATION: Randomized region.  
 148 <400> SEQUENCE: 8  
 W--> 149 ggggaagatc taaaaa nnnnnnnnnnn nnnnnnttttt aagcttgggg  
 152 <210> SEQ ID NO: 9  
 153 <211> LENGTH: 51  
 154 <212> TYPE: DNA  
 155 <213> ORGANISM: Artificial Sequence  
 157 <220> FEATURE:  
 158 <223> OTHER INFORMATION: 20 nucleotides randomized region  
 160 <220> FEATURE:  
 161 <221> NAME/KEY: misc\_feature  
 162 <222> LOCATION: (17)..(36)  
 163 <223> OTHER INFORMATION: Randomized region.  
 166 <400> SEQUENCE: 9  
 W--> 167 ggggaagatc taaaaa nnnnnnnnnnn nnnnnntttt taagcttggg g  
 170 <210> SEQ ID NO: 10  
 171 <211> LENGTH: 52  
 172 <212> TYPE: DNA  
 173 <213> ORGANISM: Artificial Sequence  
 175 <220> FEATURE:  
 176 <223> OTHER INFORMATION: 21 nucleotides randomized region  
 178 <220> FEATURE:  
 179 <221> NAME/KEY: misc\_feature  
 180 <222> LOCATION: (17)..(37)  
 181 <223> OTHER INFORMATION: Randomized region.  
 184 <400> SEQUENCE: 10  
 W--> 185 ggggaagatc taaaaa nnnnnnnnnnn nnnnnnttt ttaagcttgg gg  
 188 <210> SEQ ID NO: 11  
 189 <211> LENGTH: 15  
 190 <212> TYPE: DNA  
 191 <213> ORGANISM: Artificial Sequence  
 193 <220> FEATURE:  
 194 <223> OTHER INFORMATION: primer  
 196 <400> SEQUENCE: 11  
 197 cccaagctt aaaaa  
 200 <210> SEQ ID NO: 12  
 201 <211> LENGTH: 19  
 202 <212> TYPE: DNA  
 203 <213> ORGANISM: Artificial Sequence  
 205 <220> FEATURE:  
 206 <223> OTHER INFORMATION: 19 base pair clone  
 208 <400> SEQUENCE: 12  
 209 aaagggttta cgtggttg  
 212 <210> SEQ ID NO: 13  
 213 <211> LENGTH: 19  
 214 <212> TYPE: DNA  
 215 <213> ORGANISM: Artificial Sequence  
 217 <220> FEATURE:

Invalid response

pls explain "N" locations.

FYI:

pls explain source of genetic material.

"N" CAN

INVALID response

same error

only represent a single nucleotide

same errors

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/517,324

DATE: 12/28/2004

TIME: 11:50:05

Input Set : A:\040679.txt

Output Set: N:\CRF4\12282004\J517324.raw

```

218 <223> OTHER INFORMATION: 19 base pair clone
220 <400> SEQUENCE: 13
221 aatcgtctta tttgcatgc 19
224 <210> SEQ ID NO: 14
225 <211> LENGTH: 19
226 <212> TYPE: DNA
227 <213> ORGANISM: Artificial Sequence
229 <220> FEATURE:
230 <223> OTHER INFORMATION: 19 base pair clone
232 <400> SEQUENCE: 14
233 aattgacatg tgagcttgg 19
236 <210> SEQ ID NO: 15
237 <211> LENGTH: 19
238 <212> TYPE: DNA
239 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: 19 base pair clone
244 <400> SEQUENCE: 15
245 agtagcttgt tgaggttgg 19
248 <210> SEQ ID NO: 16
249 <211> LENGTH: 19
250 <212> TYPE: DNA
251 <213> ORGANISM: Artificial Sequence
253 <220> FEATURE:
254 <223> OTHER INFORMATION: 19 base pair clone
256 <400> SEQUENCE: 16
257 cagcatcact gtatgtgtc 19
260 <210> SEQ ID NO: 17
261 <211> LENGTH: 19
262 <212> TYPE: DNA
263 <213> ORGANISM: Artificial Sequence
265 <220> FEATURE:
266 <223> OTHER INFORMATION: 19 base pair clone
268 <400> SEQUENCE: 17
269 ctatcttcgt ggaggttgg 19
272 <210> SEQ ID NO: 18
273 <211> LENGTH: 19
274 <212> TYPE: DNA
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: 19 base pair clone
280 <400> SEQUENCE: 18
281 ctatgaaggt ggtgatgcg 19
284 <210> SEQ ID NO: 19
285 <211> LENGTH: 19
286 <212> TYPE: DNA
287 <213> ORGANISM: Artificial Sequence
289 <220> FEATURE:
290 <223> OTHER INFORMATION: 19 base pair clone

```

## RAW SEQUENCE LISTING

DATE: 12/28/2004

PATENT APPLICATION: US/10/517,324

TIME: 11:50:05

Input Set : A:\040679.txt

Output Set: N:\CRF4\12282004\J517324.raw

292 <400> SEQUENCE: 19  
293 ctttaattggt ggttgtagg 19  
296 <210> SEQ ID NO: 20  
297 <211> LENGTH: 19  
298 <212> TYPE: DNA  
299 <213> ORGANISM: Artificial Sequence  
301 <220> FEATURE:  
302 <223> OTHER INFORMATION: 19 base pair clone  
304 <400> SEQUENCE: 20  
305 tggctgtatg tgagtggct 19  
308 <210> SEQ ID NO: 21  
309 <211> LENGTH: 19  
310 <212> TYPE: DNA  
311 <213> ORGANISM: Artificial Sequence  
313 <220> FEATURE:  
314 <223> OTHER INFORMATION: 19 base pair clone  
316 <400> SEQUENCE: 21  
317 ttaatctctg gtgtcctaa 19  
320 <210> SEQ ID NO: 22  
321 <211> LENGTH: 19  
322 <212> TYPE: DNA  
323 <213> ORGANISM: Artificial Sequence  
325 <220> FEATURE:  
326 <223> OTHER INFORMATION: 19 base pair clone  
328 <400> SEQUENCE: 22  
329 ttgtagggac ttggatgat 19  
332 <210> SEQ ID NO: 23  
333 <211> LENGTH: 15  
334 <212> TYPE: DNA  
335 <213> ORGANISM: Artificial Sequence  
337 <220> FEATURE:  
338 <223> OTHER INFORMATION: Primer  
340 <400> SEQUENCE: 23  
341 aaaaattcga acccc 15  
344 <210> SEQ ID NO: 24  
345 <211> LENGTH: 50  
346 <212> TYPE: DNA  
347 <213> ORGANISM: Artificial Sequence  
349 <220> FEATURE:  
350 <223> OTHER INFORMATION: complementary to SEQ ID NO 8  
352 <220> FEATURE:  
353 <221> NAME/KEY: misc\_feature  
354 <222> LOCATION: (17)..(35)  
355 <223> OTHER INFORMATION: Randomized region.  
358 <400> SEQUENCE: 24  
W--> 359 ccccttctag attttt~~nnnn~~ nnnnnnnnnn nnnnnnaaaaaa ttcgaacccc 50  
362 <210> SEQ ID NO: 25  
363 <211> LENGTH: 35  
364 <212> TYPE: DNA

Same errors

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/517,324

DATE: 12/28/2004  
TIME: 11:50:06

Input Set : A:\040679.txt  
Output Set: N:\CRF4\12282004\J517324.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 7,9 ✓  
Seq#:8; N Pos. 17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35  
Seq#:9; N Pos. 17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36  
Seq#:10; N Pos. 17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36  
Seq#:10; N Pos. 37  
Seq#:24; N Pos. 17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35  
Seq#:25; N Pos. 11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29  
Seq#:26; N Pos. 7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25  
Seq#:34; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19  
Seq#:38; N Pos. 8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,36  
Seq#:38; N Pos. 37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54  
Seq#:39; N Pos. 4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,32,33,34  
Seq#:39; N Pos. 35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50  
Seq#:42; N Pos. 8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26  
Seq#:43; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19  
Seq#:44; N Pos. 4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:45; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19



## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/517,324

DATE: 12/28/2004

TIME: 11:50:06

Input Set : A:\040679.txt

Output Set : N:\CRF4\12282004\J517324.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:67 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0  
L:149 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0  
L:167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0  
L:185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0  
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0  
L:377 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0  
L:395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0  
L:497 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34 after pos.:0  
L:557 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38 after pos.:0  
L:583 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:0  
L:627 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42 after pos.:0  
L:645 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0  
L:663 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0  
L:681 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45 after pos.:0

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 105171324

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1      Wrapped Nucleics  
    Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2      Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.

3      Misaligned Amino  
    Numbering The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4      Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5      Variable Length Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>.<223> section that some may be missing

6      PatentIn 2.0  
    "bug" A "bug" in PatentIn version 2.0 has caused the <220>.<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>.<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>.<223> sections for Artificial or Unknown sequences.

7      Skipped Sequences  
    (OLD RULES) Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO X (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS (Do not insert any subheadings under this heading!)  
(xi) SEQUENCE DESCRIPTION SEQ ID NO X (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES" response to include the skipped sequences.

8      Skipped Sequences  
    (NEW RULES) Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
<210> sequence id number  
<400> sequence id number  
000

9      Use of n's or Xaa's  
    (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>.<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10      Invalid <21>  
    Response Per 1.823 of Sequence Rules, the only valid <21> responses are Unknown, Artificial Sequence, or scientific name (Genus/species). <220>.<223> section is required when <21> response is Unknown, Artificial Sequence, or scientific name.

11      Use of <220>

Use of <220> to <223> is MANDATORY if <21> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12      PatentIn 2.0  
    "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13      Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid